Backlog in the Pleiades PBS Queues

Category: Under Investigation

Problem: Pleiades users have experienced longer wait times for PBS jobs due to heavy loads on the queues served by pbspl1.

Status: Under Investigation

Actions

Updated 04.02.12 - On March 21st, 24 Harpertown racks were disconnected and removed from Pleiades to provide the space needed for the addition of Sandy Bridge racks. The new racks will more than triple the peak performance of the racks that were removed. The new Sandy Bridge racks are expected to become available by late April.

Updated 02.10.12 - To increase compute capacity, the HECC project has ordered 1,728 new SGI ICE X nodes with next-generation processors. Installation will begin at the end of March 2012. When the new nodes are in production, they will provide more than 3 times the peak performance of the 1728 Harpertown nodes which they are replacing. We expect that the new processors will provide a 6-fold increase in productivity on the HECC workload when compared to the Harpertown processors being removed.

Tips

- If you are doing code development and your PBS job requires less than 2 hours of walltime, submitting it to the **devel** queue should provide faster turnaround; see Pleiades devel Queue for more information
- The 64 "bigmem" nodes (rack 32, Harpertown processors) are usually lightly utilized. If your job needs 8-16 GB of memory per node, you can request the bigmem nodes, which may provide faster turnaround:
    - Add **:bigmem=true** to your resource request
    - See the first tip listed in the article How to Get More Memory for your Job
- If your PBS job can run on any of the current processor types (Harpertown, Nehalem, and Westmere), try submitting your job to the processor type with the lowest demand; we recommend the following tools:
    - Use **/u/scicon/tools/bin/qs** to get information on the number of used and unused nodes for each processor type
    - Use **/u/scicon/tools/bin/node_stats.sh** to find out the number of nodes for each processor type currently requested by queued jobs

- If you are submitting serial jobs (in which each PBS job requests 1 node and uses only 1 core in the node), you may be able to package multiple serial runs into a single PBS job, in order to use more cores in the node. This reduces the burden of the PBS scheduler by decreasing the total number of PBS jobs, and frees up some resources for other users. Use one of the two approaches listed in the following articles:
    - Running Multiple Serial Jobs to Reduce Walltime
    - Using GNU Parallel to Package Multiple Jobs in a Single PBS Job
- If some of your jobs are not time sensitive, consider waiting until the backlog decreases before submitting them; if any jobs are optional, consider submitting them later
- If you have an urgent deadline to meet and must get some of your PBS jobs completed, you may need to coordinate with other users of your Mission Directorate; if needed, contact your mission directorate POC and/or send an email to support@nas.nasa.gov

Background

The demand for Pleiades compute resources for PBS jobs is very heavy, and users' jobs have been experiencing longer wait times since late December 2011. The average wait time is likely to increase for a short time when the HECC project begins the process of retiring 24 Harpertown racks to make room for the new racks.

Article ID: 304
Last updated: 13 Aug, 2012
Updates on Issues -> Under Investigation -> Backlog in the Pleiades PBS Queues
http://www.nas.nasa.gov/hecc/support/kb/entry/304/?ajax=1